# The Problem of Attitude Prediction Based on Sentiment Implicatures

Klenner, Manfred



Originally published at:
Klenner, Manfred (2016). The Problem of Attitude Prediction Based on Sentiment Implicatures. In: Computational modeling of attitudes 4th Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2016), New York, 9 July 2016. ResearchBib, 6-11.

# The Problem of Attitude Prediction Based on Sentiment Implicatures

**Manfred Klenner**
Computational Linguistics
University of Zurich
Switzerland
klenner@cl.uzh.ch

## Abstract

We introduce a novel rule-based approach to attitude prediction based on sentiment implicatures. It is based on a verb lexicon specifying so-called polarity frames where the semantic roles of the verb might cast positive or negative polar effects on the argument fillers. Verbs that subcategorize for clausal complements are further classified according to their signature, i.e., whether they assert factuality on the subclause. An empirical evaluation revealed that (naive) human annotators determine polar effects rather selectively. This is in contrast to the rigorous behaviour of our current rules. This raises the need for a more sophisticated theory of attitude prediction.

## 1 Introduction

Recently, attitude prediction has raised some interest in the field of sentiment analysis (e.g. [Deng and Wiebe, 2015]). It strives to determine what the overt or (more interesting) hidden attitudes among the participants of a given sentence are. Verbs seems to be crucial here. For instance, a verb might cast a positive or negative effect on its direct object as in "A helps B". It is good for B to be supported. But is this true in any case and what kind of positive effect is it: is it intrinsically or extrinsically good. Given the polar effects of a verb, the attitudes among the participants can be captured with rules like: if A is (lets say) the actor of a situation described by a verb that casts a positive effect on its (lets say) theme B, then A is positive towards B. If A and B are humans or organisations etc. we would say that they are proponents. More complicated cases arise if we consider subclause embeddings as in "C criticizes that A helps B". Now, A is still a proponent of B, but what about C? Obviously, C is against A and B. Again, we need to know that "criticize" casts a negative effect on its clausal complement, and we need general rules that tell us that an agent C is an opponent of an agent B if C is against a situation (help) that casts a positive effect on B. Similarily, C is an opponent of an agent A who is in charge of the positive effect on B.

Can we infer anything if negation is present: C has not criticized that A has not helped B? This is no longer only a semantic question, but a pragmatic one as well. It could be the case that the writer of such a sentence simply negates that such an event has taken place at all. Then nothing follows, it is a kind of denial that happens: C might have been unaware of any situation where B has needed help that A was not able or willing to give. But if the sentence is meant as a reproach, namely that C should have criticized ("that .."), then again attitudes become apparent. First of all, the writer is against C (and A) and pro B. C, on the other hand, must be somehow against B, because a negative effect on B was not worth a (negative) comment in the eyes of C. What about C and A? Is C an proponent a A (well, he has not criticized the omision of help of A)? In a sense, C is pro A, but not actively, more implicitely. But is this a safe inference, do humans agree on it?

In this paper, we introduce a rule-based model for German able to draw these kind of inferences. Our model not only predicts attitudes, but also determines who benefits (or suffers) from the situation described by a sentence. This is in contrast to existing work which only deals with attitudes. Moreover, our approach is the first one that takes factuality into account. Factuality is crucial for a certain kind of inferences. The condition under with factuality can be infered are verb-specific and depend on the affirmative status of the sentences.

In an empirical evaluation, we found that humans do not agree to the extent we expected and, thus, system performance was difficult to evaluate. In this paper, we discuss the reason for this and its implications.

## 2 Verb Polarity Frames

A verb polarity frame is a subcategorization frame of a verb where some verb roles cast polar effects given that a sentence with the verb is factual and affirmative (not negated). If A criticizes B, this is negative for B. If A wins the competition, this is positive for A, and finally, if A hurts B, this is negative *of* A and negative *for* B. These effects are associated with argument positions of the verb frame. We distinguish $a_1$, $a_2$, $a_3$ and $a_4$, capturing the (logical) subject, the object (direct and indirect), the prepositional complement and the subclause, respectively. The logical subject of a verb instance might be given by the direct object of the matrix clause it is embedded in (if the matrix verb is an object control verb), we thus prefer to talk of argument positions instead of grammatical roles. Note that in our current model "(The famous) Bocuse cooks" and "The rice cooks" would have the same argument structure. Selectional restrictions make the difference here. Also verb ambiguity is something that needs to be coped with outside our model (e.g. in the preprocessing phase).

In order to motivate a further distinction, let us have a look at a candidate for a general inference rule. Given that A criticizes that B injures C. If A disapproves (here "criticize") an event (injury) that is negative *for* C, then we are inclined to believe that A is a proponent of C, we say, A is pro C. Also, if B is the agent *of* that event, then probably he is an opponent of A, we say: A is con B. It turns out that the distinction between *of-roles* and *for-roles* is helpful for specifying general inference rules. Instead of assigning positive or negative effects to a verb argument, we specify whether it is a positive or negative of-role or for-role. Table 1 shows three frames of the verb "to criticize".

| Role | $F_1$ | $P_1$ | $F_2$ | $P_2$ | $F_3$ | $P_3$ |
|---|---|---|---|---|---|---|
| $r_1$ | $a_1$ | of | $a_1$ | of | $a_1$ | of |
| $r_2$ | $a_2$ | nfor | $a_2$ | nfor | $a_4$ | neff |
| $r_3$ | $a_4$ | neff | - | - | - | - |

Table 1: Frames of "to criticize"

Frame $F_1$ comprises three roles, r1 to r3 (A criticizes B for C). Role $r_3$ is realized as $a_4$, e.g. a clause level argument in the form of a gerund. We distinguish the following polarity roles (P): pfor, nfor, pof, nof capturing positive and negative for- and of-roles (of and for denote neutral cases). Clause level effects are specified as negative (neff) or positive (peff). Polarity roles like pfor are generalizations, e.g. $a_1$ and $a_2$ can both realize an for-role given particular verbs.

To summarize, verbs are characterized by their arguments $a_i$. Different syntatic constellations might give rise to the instantiation of a particular $a_i$ (passive voice, control verbs, etc.). Within a frame, an $a_i$ is associated with a polarity role (pfor, etc.). Inference rules rely on the status of an argument as a polarity role, independent of the concrete argument type (e.g. $a_1$ or $a_2$).

The current German verb lexicon comprises 330 different verbs resulting in 680 polarity frames. About 80 verbs subcategorize for a complement clause. Although the model is for German, all examples are given in English.

## 3 The Role of Factuality

A major claim of this paper is that factuality plays a crucial role in the determination of attitudes and also for polar effects posed by particular verbs. Consider for illustration the simple case of a positive effect that the subject receives given the sentence *The president wins the election*. Various approaches from the literature rather focus on the attitudes between the participants or even those of the author of a text towards the participants (e.g. [Deng and Wiebe, 2014]). They do not care for factuality. But in these scenarios, factuality plays a role, although not in each and every constellation. Attitudes that stem from intra-clausal dependencies require factuality. For instance, in *The president criticizes the minister* a negative attitude of the president towards the minister holds if the sentence is factual. If the sentence is embedded in a verb casting non-factuality like *to hope*, this no longer is true (cf. *The vice president hopes that the president criticizes the minister*). This sentence, however, also shows that predictions between an actor of a (factual) matrix clause and those of a (non-factual) subclause are possible. Here the vice president has negative attitude towards the minister although the subclause is non-factual. Factuality is crucial, but there are also constellations, where it neither licenses nor prevents attitude predictions.

## 4 Verb Signatures

Verbs that subcategorize for a clausal complement are further specified for (non-)factuality of the clausal complement. Factuality means that the situation described in the subclause is meant (by the writer) to be true (to hold). We follow the work of [Karttunen, 2012] who distinguishes factive, non-factive and implicative verbs. Factuality of the subclause depends on the matrix verb signature and the presence or absence of negation in the matrix clause. Factive verbs such as "to regret" cast factuality on their subclause, irrespective of whether the main clause is negated or not. If A regrets that COMP, then COMP (the subclause) is true in the sense that the speaker believes (or a least asserts) COMP to be true. The same holds for A does NOT regret that COMP (factuality here is constant under negation, thus factuality is a presuppositon of factive verbs). Note that factive verbs need not to be factual in order to cast factuality. If A hopes that B regrets (non-factual) to bother C, then the bothering event is factual. Sublcauses of non-factive verbs, on the other hand, are never meant to be true (e.g. "to pretend", "to hope").

Then there are verbs called implicatives that cast a mixture of factuality and non-factuality. Two-way implicatives like "to forget to" have non-factual subclauses in an affirmative (not negated) use, but factual subclauses if negated. One-way implicatives only give rise to factuality in either the affirmative (like "to force") or negated matrix verb contexts (like "to refuse"). For instance, if A forces B to lie, B lies. If A does not force B to lie, then B might lie as well, we just cannot tell. Table 2 summarizes the signatures, introduces the concept labels (e.g. AF) we use to represent them and gives example verbs.

| label | explanation | matrix verb |
|---|---|---|
| F | factual in any case | to regret |
| NF | non-factual in any case | to hope |
| AF | factual, if affirmative | to force |
| ANF | non-factual, if affirmative | to forget |
| NaF | factual, if non-affirmative | to forget |
| NaNF | non-factual, if non-affirmative | to manage |
| NaO | true or false, if non-affirmative | to help |

Table 2: (Non-)Factuality of Subclauses

We found this information to be crucial for inferences. Non-factuality blocks some, but not all inferences. Take: "A hopes that B wins". The subclause is non-factual, so B does not receive a positive effect (he is not a beneficiary): this inference is blocked. However, the attitude of the of-role bearer of the (factual) matrix sentence (A) towards the bearer of the for-role (B) of the (non-factual) embedded verb still holds (a positive relationship): it is not blocked.

Relationship inferences *within* a non-factual clause, however, are blocked. For instance, if A hopes that B loves C, the inference that B has a positive attitude towards C is blocked.

| | of | for | sc-eff | aff | neg |
|---|---|---|---|---|---|
| to criticize | of | - | neff | AF | NaF |
| to approve | of | - | peff | AF | NaF |
| to help | pof | - | peff | AF | NaO |
| to help | pof | pfor | - | | |
| to injure | nof | nfor | - | | |
| to survive | - | pfor | - | | |

Table 3: Polarity Frames

In Table 3 we give the polarity frames including signatures of some verbs. A hyphen indicates that the role is not part of the verb frame in question, column sc-eff means subclause effect. The last two columns relate to the verb signatures as introduced in Table 2, the penultimate column reports the restriction if the matrix verb is aff(irmative) and the last column if it is neg(ated). For example, the subclause of "to help" (line 3) is factual if the help sentence is affirmative (AF), but its truth value is unspecified (NaO) if negated.

## 5 Rule-based Sentiment Inference

Our rule-based approach is realized in Prolog, the rule interpreter is implemented as a meta interpreter. Starting from predicate argument structures, the algorithm determines factuality from outside-in (from the matrix verb to the innermost verb), then at the innermost level all rules are applied. The algorithm keeps doing this (rule application) at each recursion level gathering all the assertions so far derived until the outermost level is reached (again).

The syntax of the rules we propose is quite simple. We use the Prolog notation here, since all rules are basic horn clauses with equivalents in predicate logic. The only notational convention to mention is that the head of such a rule comes first, separated from the body by ":-" (the implication operator turned to its left ←). Variables are written in caps, a comma (,) means conjunction, a semicolon (;) indicates disjunction, \+ means negation (as failure). Take the following definition of beneficiary: a beneficiary X is someone who takes the pfor role in an affirmative and factual sentence I:

```
b1 beneficiary(X,I):- fact(I),aff(I),pfor(I,X).
```

The equivalent predicate logic formula is:
$\forall I, X$: pfor(I,X)$\wedge$ aff(I)$\wedge$ fact(I) $\rightarrow$ beneficiary(X,I)

In order to introduce our scheme, we go through the following (hypothetical) sentence, the input structures are given in Table 4.

S: *The minister has criticized that the EU has helped Greece to survive*.

These instantiations are based on the polarity frames of the verbs and the dependency tree of the sentence. Since no negation is present, it holds that aff(criticize), aff(help),

| # | input predicates | |
|---|---|---|
| 1 | of-role(criticize,minister) | neff(criticize,help) |
| 2 | pof(help,EU) | pfor(help,Greece) |
| 3 | peff(help,survive) | pfor(survive,Greece) |
| 4 | aff(criticize) | aff(help) |
| 5 | aff(survive) | fact(criticize) |

Table 4: Input Representation

aff(survive) (line 4 and 5), where aff means affirmative use. The matrix clause (since no modal is present) is factual (line 5), i.e., fact(criticize). Note that pfor(help,Greece) just means that Greece occupies a particular polar role. Whether Greece actually gets a positive effect depends on the factuality as determined by the matrix verb and its affirmative status (and also the affirmative status of the complement verb itself).

The factuality of complement clause I is determined by the following rules (and based an the verb signatures, the predicate is vsig(Verb,Signature)):

```
f1 fact(I) :-
     eff_role(MCI,I),vsig(MCI,AF),aff(MCI).
f2 fact(I) :-
     eff_role(MCI,I),vsig(MCI,NaF),\+aff(MCI).
f3 fact(I) :- eff_role(MCI,I),vsig(MCI,F).
where:
eff_role(MCI,I):- neff(MCI,I);peff(MCI,I).
```

A sentence I (I is the finite verb) is factual if it is embedded into a matrix verb MCI and MCI is affirmative and has verb signature AF (asserted factual), if it in negated (\+aff) and bears signature NaF (not asserted factual) or if it is a factive verb (f3). A factive verb like "regret" casts factuality independent of its own factuality status: If A hopes that B regrets (non-factual) to bother C, then the bothering event is factual. If A hopes that B helps (non-factual) C to survive, although "help" is asserted and is of type AF (i.e. casts factuality if asserted), in the context of "hope" (a ANF) this is blocked: the survive event is non-factual.

Since eff_role(criticize,help), vsig(criticize,AF) and aff(criticize) satisfy f1, it follows that fact(help). Similarly, fact(survive) can be derived. Now can see that Greece is a beneficiary (see the definition b1 of beneficiary) since (as derived by f1) fact(help) and (from Table 4) aff(help) and pfor(Greece, help), which gives i1 from Table 5.

Before reading the further outline of our rule component, the reader is invited to verify that the following inferences drawn from the example sentence S are in line with his/her intuition (i4 and i6 needs further explanation, though):

| # | inference | rule |
|---|---|---|
| i1 | beneficiary(Greece) | b1 |
| i2 | pro(EU,Greece) | r1 |
| i3 | con(minister,EU) | r2 |
| i4 | disapprove(minister,survive) | r3 |
| i5 | con(minister,Greece) | r4 |
| i6 | con(EU,minister) | r6 |

Table 5: Inferences

We now introduce the rules needed to understand our example. The main goal is to find out, whether A is for B, which we model with the property pro; or whether A is against B, here con is used. A verb might (directly) reveal the relation between the participants within the same clause: if A helps B, then A is pro B. If A criticizes B, then A is con(tra) B (at least in a certain - the given - context, not necessarily in a fundamental, irreconcilable way). Provided, of course, the situation is factual.

```
r1 pro(X,Y) :-
 of_role(I,X),fact(I),aff(I),pfor(I,Y).

where: of_role(I,X) :- pof(I,X);nof(I,X).
```

Rule r1 states: An actor X (the of-role) is pro Y if in a single factual, affirmative sentence I, Y is the filler of the pfor role (i2 from Table 5 follows): pof(EU,Greece).

If a sentence I embeds a sentence I2, then rules like the following are in charge:

```
r2 con(X,Y) :- aff(I),fact(I),neff(I,I2),
      aff(I2),of_role(I,X),of_role(I2,Y).
```

According to r2, an affirmative and factual matrix clause I that embeds an affirmative sublause I2 (factuality of I2 is irrelevant) bearing a negative effect ( neff) gives rise to a con relation between the of-role of the matrix clause and the of-role of the subclause (see i3 from Table 5): con(minister, EU).

More complicated scenarios arise in the case of multiple embeddings. According to Table 3, "to criticize" has a neff role while "to help" has a peff role. If someone A criticizes that someone B helps somebody C to achieve something D (D=survive), then, obviously, A disapproves with D. That is, a neff on a peff gives disapprove, see rule r3.

```
r3 disapprove(X,I3) :-
     neff(I,I2),peff(I2,I3),
     aff(I),fact(I),aff(I2),of_role(I,X).
```

The matrix clause must be factual: if A (just) *might* criticize that COMP, nothing can be inferred about A's (dis-)approval regarding COMP (and COMP of COMP). Rule r3 triggers and produces i4 from Table 5:disapprove(minister,survive).

The next rule describes how disapprove propagates to a con relation (factuality is irrelevant here).

```
r4 con(X,Y) :-
      disapprove(X,I),aff(I),pfor(I,Y).
```

If someone disapproves an affirmative situation that is positive (pfor) for someone, then he is against this person. Rule r4 produces i5 from Table 5: con(minister,Greece).

Another way to come to the same conclusion (i5) is to combine pro and con relations. If A is con B and B is pro C then A is con C (rule r5):

```
r5 con(X,Z) :- con(X,Y),pro(X,Z).
r6 con(X,Y) :- pro(X,Z),con(Y,Z).
```

There are some cases which we cannot describe using con/pro inferences but only with intermediate (dis-)approve derivations. There are also inferences that need con/pro. If, for instance, A is pro B and C is con B then we might be allowed to guess that A is con B. In our example it follows EU con minister (rule r6), see i6 from Table 5. Note that these transitively given pro and con relations must not be taken too fundamentally. If (rule r6) A admires B while C finds B boring, both, A and C are opponents, but only conditional on B, so to speak. In general, pros and cons can only deliver situation-specific attitudes.

## 6 A Corpus for Sentiment Inference

No German gold standard exists for our inference task. We thus have started to create such a corpus. Before starting a more ambitious initiative with real sentences, we first wanted to explore how our model performs given rather complicated sentences. However, a small number of real sentences were considered as well. We started with 100 made-up (corpus F) and 40 real sentences (corpus R). Corpus F, the constructed sentence corpus, is meant to represent the phenomena we are dealing with in a very condensed way. That is we created complex sentences with up to four levels of embedding, where each subclause covers a verb from our lexicon. We also created versions of some of the sentences where negation is distributed over all (sub)clauses exhaustively according to the possible permutations. In order to avoid a bias towards verb-specific phenomena, 50 different verbs were used.

Another question is whether we can clearly define the task we intend to solve. What are the annotation guidelines? Our two annotators are given the following instructions: Take the sentence to be true, then

- for all pairs of entities from the sentence, determine whether there is an attitude of one entity towards the other entity (and vice versa) - and whether it is positive or negative.
- for all entities from the sentence, determine whether this entity can be seen as benefitting or suffering (in the broadest sense) from the situation described.

We also told them that the attitudes in question are to be understood as situation-specific and need not necessarily hold in general. We deliberately avoided to give them any background information concerning our model and the details of our verb resource.

The other 40 sentences, corpus B, were taken randomly from the DeWac corpus [Baroni *et al.*, 2009]. Again, only sentences were considered where at least two verbs from our lexicon are forming a complex clause.

## 7 Evaluation

The output of the annotation efforts in terms of the confusion matrices was a bit of a surprise. We expected a higher agreement, since we thought the task at hand was straightforward and the annotation guidelines were precise. However, the annotation task of the sometimes rather complex constructions does not seem to be that simple. The task was to annotate pro, con, sufferer and beneficiary cases. Table 6 shows the confusion matrix of the pro/con annotation (annotator A and B).

We can see that the main problem is the none class, where one annotator has chosen to annotate a relation while the

| | | B pro | B con | B none | |
|---|---|---|---|---|---|
| A | pro | 42 | 3 | 16 | 61 |
| A | con | 1 | 126 | 28 | 155 |
| A | none | 13 | 37 | 0 | 50 |
| | | 56 | 166 | 44 | 266 |

Table 6: Confusion Matrix: pro and con

other one did not. For instance, there are 37 cases where annotator B thought a con relation should hold, while annotator A did not.

The situation is similar with beneficiary/sufferer classifications (see Table 7).

| | | B benef. | B sufferer | B none | |
|---|---|---|---|---|---|
| A | benef. | 30 | 1 | 11 | 42 |
| A | sufferer | 2 | 57 | 36 | 95 |
| A | none | 4 | 17 | 0 | 21 |
| | | 36 | 75 | 47 | 158 |

Table 7: Confusion Matrix: benef. and sufferer

These results show that the task is more difficult than expected. We have avoided to train the annotators exhaustively in advance since we wanted the results to be free of any decisions stemming from model-conform considerations.

We compared the output of our system with the annotations produced by the two annotators. Although we have a running pipeline, in this study, the system was provided with perfect predicate argument structures to enable a fair comparison of the performance of the rules to human performance (since humans are not impaired by noisy syntactic structures).

If we take the annotations of A as gold (see Table 8 and 9), then the precision of our system for pro and con detection is 58% while recall is 63%. If we evaluate annotator B against annotator A taken as gold, precision is 77% and recall 78%.

| | | System pro | System con | System none | |
|---|---|---|---|---|---|
| A | pro | 39 | 4 | 19 | 62 |
| A | con | 7 | 101 | 53 | 161 |
| A | none | 47 | 43 | 0 | 90 |
| | | 93 | 148 | 72 | 313 |

Table 8: Confusion Matrix: Annotator vs System

For the detection of beneficiariy and sufferer, precision is 64% and recall 76% (see Table 9). If we evaluate annotator B against A taken as gold, precision is 81% and recall 64%.

The overall picture is not very clear. While the f-measure for pro/con 77.5% (B) versus 60.4% (system) makes a big difference, the one for sufferer/benef., 71.2% (B) versus 69.5% (system), does not. System performance is behind human agreement, which is on the other hand is far from being perfect.

| | | System benef. | System sufferer | System none | |
|---|---|---|---|---|---|
| A | benef. | 35 | 0 | 7 | 42 |
| A | sufferer | 2 | 72 | 24 | 98 |
| A | none | 10 | 50 | 0 | 60 |
| | | 47 | 122 | 31 | 200 |

Table 9: Confusion Matrix: Annotator vs System

## 8 Error Analysis

There are various reasons for the relatively poor agreement among human annotators. First of all, the pragmatic status of a negated sentence (matrix verb) might be unclear. As previously mentioned, if writer reports that C has not criticized that COMP, then either C has intentionally not criticized it or he was not aware of COMP at all. Only in the first case (a reproach), we are allowed to infer attitudes of C towards the participants of COMP. The second case is a plain denial.. This problem, however, depends on the verb. If the writer reports that A has not managed (instead of criticize) to COMP, then the denial reading is not valid, since "not to manage" presupposes an attempt to do so. In this case, sentiment inferences are licensed. This suggests that sometimes a pragmatic decision has be taken (criticize) and sometimes the presence of the verb is sufficient. Our current system has no means to draw pragmatic inferences.

Also verb semantics beyond what we have captured via our polar effects is crucial. In such cases, factuality alone no longer might license a pro or a con relation. If A *forces* B to help C, then it cannot be safely inferred that B is pro C. The semantics of "force" rather blocks positive or negative attitudes of the embedded participants (even if factuality is given). However, polar effects are not blocked, C still is a beneficiary.

Also common sense reasoning might contribute to attitude prediction. If A is an unagreeable person (e.g. a terrorist) then it no longer holds that if (such an) A supports B this should count as a positive effect for B. That is, if B is con (such an) A (like most people are) then an inference via the verb "support" that A is pro B produces a conflict that should prevent the assertion of a positive effect on B. However, if it is the other way round, e.g. that A is an agreeable person and A criticizes B, where a negative effect should trigger, and given that B is pro A, then the negative effect on B actually should occur (i.e. the A con B relation does not produce any complications). Our model is not fine-grained enough for these cases, at the moment.

The inclination to assign polar effects in the presence of negation seems to be gradual and dependent on the verb. If A has not survived (some catastrophe), a negative effect on A is uncontroversial, but if A has not won a bar of chocolate, a negative effect is less likely to be annotated. Priming effects might have played a role in our annotation experiments (negated survive followed by negated win).

From the sentence *A fears that he misses the train* a negative effect on A could be derived, although *miss the train* is neither factual nor a very drastic catastrophe. The negative effect could not come from "miss" since it is not factual, it must come from "fear". This is only possible if the polarity frame of "fear" specifies a nfor role. However, in *A fears that B misses the train*, a negative effect on A is harder to claim (it is rather week, one could argue). Should A actually occupy a nfor role of "fear" in this case? Or should this only be allowed if the subject of fear and the subject of the embedded verb are identical? Our current model does not make this distinction.

Cognitive complexity is another problem, e.g. massive negation. Take *A criticized that B has not convinced C not to offend D*. First of all, the verb"convince" is special, since if both, convince and its subclause are negated, the subclause event turns out to have happened. So D was offended. If we replace "convince" through "help" (B has not helped C not to offend D), D was not offended by anyone. This kind of behaviour increases cognitive complexity and might be the reason for wrong annotations. Our current rule component is not yet able to cope with these special kinds of verbs (like convince). In case of a double negation (in the context of such a verb), a wrong prediction results, thus.

There are also a number of very specical cases that we cannot help but to report. Take a real sentences from our corpus: *The strategic maneuver of the SP* (a political party)*to recommend Hollenstein in order to prevent Bortoluzzi, was successful*. One annotator claimed: con(Hollenstein, Bortoluzzi) obviously on the grounds that they are rivials in a election, but this rivalery is mediated only by a third party (the SP).

Given A approves that B lies. One annotator claimed a negative effect on B (since B is bound to do something immoral), the other one a positive one (from approve). Our system assigns a pro relation (A pro B) and a nof relation, a negative of-role (the annotators were instructed to ignore these cases - actually, they are too rare in our corpus). Disagreement is maximized in this case, we would say that the system's inferences are right.

In this section, we have tried to find the reasons for the low agreement among human annotators. We have identified a number of problems that also indicate that our rule component needs to be refined. The current version is slightly behind (the not very striking) human performance, one reason for this is that our rules are (in part) not fine-grained enough.

In order to improve annotator agreement, we have to improve the annotation guidelines. In order to improve system performance, we have to refine our model.

## 9 Related Work

An early rule-based approach to sentiment inference is [Neviarouskaya *et al.*, 2009]. Each verb instantiation is described from an internal and an external perspective. For example, "to admire a mafia leader" is classified as affective positive (the subject's attitude towards the direct object) given the internal perspective while it is (as a whole) negative externally. Factuality and subclause embedding do not play any role in their work. The same is true for [Reschke and Anand, 2011]. They capture the polarity of a verb frame instantiation as a function of the polarity of the verb's roles - we, instead, do not know in advance, but intend to infer the (contextual) polarity of the roles.

Recently, [Deng and Wiebe, 2014] and [Deng and Wiebe, 2015] have introduced an advanced conceptual framework for inferring (sentiment) implicatures. Their work is most similar to our approach. Various model versions exist, the most recent one [Deng and Wiebe, 2015] also copes with event-level sentiment inference, which brings it even closer to our model. Probabilistic Soft Logic is used for the definition of the model and for drawing inferences. The goal of the systems is to detect pairs of entities that are in a PosPair or NegPair relation. This is similar to our pro and con relations. However, as we mentioned before, factuality does not play a role in their framework, while we believe it is crucial for some inferences.

## 10 Conclusion

We have introduced a novel model for the propagation of polar effects (benefits, suffers) and attitude prediction (pro, con relations). For the first time, factuality is regarded as crucial in the context of attitude prediction. Our rule component clearly indicates that factuality determination is needed in order to properly draw these inferences.

Our empirical evaluation shows that the task at hand is not trivial. The agreement among humans was far from being perfect and - as a consequence - our system was hard to evaluate. The reasons can be found in the complexity of the task - as we have shown in section 8. Most of the problems can be solved within our choosen rule-based framework - this is future work. But there is also a need to understand how humans came to their decicisions and whether a more psychological model of attitude predication was benefical.

## References


[Baroni *et al.*, 2009] M. Baroni, S. Bernardini, A. Ferraresi, and E. Zanchetta. The WaCky Wide Web: A collection of very large linguistically processed Web-crawled corpora. *Language Resources and Evaluation*, pages 209–226, 2009.

[Deng and Wiebe, 2014] Lingjia Deng and Janyce Wiebe. Sentiment propagation via implicature constraints. *Meeting of the European Chapter of the Association for Computational Linguistics (EACL-2014)*, 2014.

[Deng and Wiebe, 2015] Lingjia Deng and Janyce Wiebe. Joint prediction for entity/event-level sentiment analysis using probabilistic soft logic models. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015, Lisbon, Portugal, September 17-21, 2015*, pages 179–189, 2015.

[Karttunen, 2012] Lauri Karttunen. Simple and phrasal implicatives. In *Proc. of the First Joint Conference on Lexical and Computational Semantics*, SemEval '12. Association for Computational Linguistics, 2012.

[Neviarouskaya *et al.*, 2009] Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka. Semantically distinct verb classes involved in sentiment analysis. In *IADIS AC (1)*, pages 27–35, 2009.

[Reschke and Anand, 2011] Kevin Reschke and Pranav Anand. Extracting contextual evaluativity. In *Proceedings of the Ninth International Conference on Computational Semantics*, pages 370–374, 2011.